Electronic Supplementary Materials

for Delayed and inefficient local immune pressure explains persistent oral cytomegalovirus shedding during primary infection in infants

Contents

Supplementary Methods

Selection of subjects.

A total of 22 CMV infected infants were identified. Two were congenital and were excluded. We excluded another 5 postnatal infections that had less than 125 days of data after acquisition, because it was not possible to identify trends (included infants had at least 244 days of follow-up) due to slow viral dynamics. A sixth infant who exhibited prolonged shedding but also signs of resolution and irregular patterns was also excluded (Infant O in Supplementary Figure S1).

Oral shedding episodes occasionally preceded fulfillment of the qPCR criteria used defining for primary infection in the original cohort study. Of the 759 swabs analyzed, there were 4 negative swabs (two for Subject B). Given the trends of the data, we assumed that these swabs were not true negatives but were either censored (below detection) or swabbing errors, and excluded them as missing data.

Calculating rebound with cubic regression.

To explore the non-linear clearance dynamics, we fit a cubic regression model to the viral clearance phase of the data for each participant (i):

$$V_i = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \epsilon_i$$

Using the fits from this regression, we can identify key dynamic points in the data. If the roots for the first derivative exist, then these critical time points correspond the local minimum and maximum of the cubic function:

$$t_{\text{critical}} = \frac{-\hat{\beta}_2 \pm \sqrt{\hat{\beta}_2{}^2 - 3\hat{\beta}_3\hat{\beta}_1}}{3\hat{\beta}_3}$$

Specifically, if the second derivative

$$\frac{\mathrm{d}^2\hat{V}_i}{\mathrm{d}t^2} = 2\hat{\beta}_2 + (3)(2)\hat{\beta}_3 t$$

is positive at a critical point then that is a local minimum. If this local minimum exists

and occurred during the time-range of the fitted clearance phase, then we called that the rebound time (denoted ** in Table S5). If the rebound time was near the end of the clearance phase then it is difficult to establish whether that is deceleration or a rebound point without extrapolating outside of the data. We denoted this case as a * in Table S5.

For the cases where the roots of the first derivative do not exist, we studied the inflection point:

$$t_{\text{inflection}} = \frac{-\hat{\beta}_2}{3\hat{\beta}_3}$$

to determine if there was a tapering. In these data we observed concavity after the inflection point and corresponding acceleration of clearance as time moves past the inflection point. Thus, in cases where the inflection point occurs long before the final observation point, there is no evidence of tapering and rebound. For two cases, the inflection point was near the final observation point with no minimum. Because the cubic clearance slope is slowing from the convex phase through the inflection point into the concave phase, this may correspond to deceleration within the time-frame of the data near the end of observation. This was confirmed visually and these two episodes were denoted * for deceleration only in Table S5.

Target-cell limitation mathematical model.

The target-cell limitation mathematical model attempts to capture the dynamics of within-host infection in the absence of immunity.

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{d}t} &= \lambda - \mu S - \beta SV \\
\frac{\mathrm{d}L}{\mathrm{d}t} &= \beta SV - \alpha L \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= \alpha L - \delta I \\
\frac{\mathrm{d}V}{\mathrm{d}t} &= pI - cV
\end{aligned} \tag{S1}$$

In this model, there are 4 population compartments, and 7 rate parameters. The compartments are susceptible cells (S), infected cells with non-replicating CMV (L), infected cells with replicating CMV (I), and free virus (V). For the susceptible population, λ and μ describe the normal turnover of cells; β represents viral infectivity rate per susceptible cell; α is the infected cell latency rate (time before virus begins replicating); δ represents the death rate of CMV-infected cells with replicating virus; p is the viral burst rate per infected cell; and c is the natural clearance of virus. Only the viral infectivity parameter (β) and infection time (between the last negative and first positive swab) were unknown and fit from the data. The parameter values for the model are shown in Supplementary Table S1.

Cytolytic immune response mathematical model.

We used the following mathematical model to capture cell-mediated immune activation:

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{d}t} &= \lambda - \mu S - \beta SV \\
\frac{\mathrm{d}L}{\mathrm{d}t} &= \beta SV - \alpha L \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= \alpha L - \delta I - kIE \\
\frac{\mathrm{d}V}{\mathrm{d}t} &= pI - cV \\
\frac{\mathrm{d}E}{\mathrm{d}t} &= \theta \frac{I}{I + K_I} - \gamma E.
\end{aligned} \tag{S2}$$

This model includes four additional parameters: the immune removal rate per infected cell, k; the maximum growth rate of immune effectors, θ; the infected cell level where immune effectors growth is 50% maximal, K_I; and the death rate of immune effectors, γ. We fixed k (1; 2) and fit the remaining three parameters from the data (Supplementary Table S2). Our goal is to capture broad aspects of immune pressure (such as timing and extent) and so we are not focused on capturing the individual values of a given parameter or the functional form of the immune response.

Alternative viral-mediated model

We also fit an immune model where virus activated and was targeted by the immune system:

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{d}t} &= \lambda - \mu S - \beta SV \\
\frac{\mathrm{d}L}{\mathrm{d}t} &= \beta SV - \alpha L \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= \alpha L - \delta I \\
\frac{\mathrm{d}V}{\mathrm{d}t} &= pI - cV - kVE \\
\frac{\mathrm{d}E}{\mathrm{d}t} &= \theta \frac{V}{V + K_v} - \gamma E.
\end{aligned}$$

This model used the parameter values specified in Tables S1 - S2. Model fits were comparable to that of the infected cell targeted model presented in the main text (data not shown). Because of the linear relationship between I and V, distinguishing the models requires data that is currently unavailable: either cell population measurements from each study participant or stricter biological boundaries on the immune parameters. We chose to present only the infected cell mediated immune model because of known etiology of immune action against CMV (3) and for succinctness of message.

R_0 and effective reproduction number, R.

The basic reproduction number, R_0, the expected amount of cells that a single infected cell infects in a fully susceptible population, was calculated from the target-cell limitation model (Equation S1) as follows:

$$R_0 = \frac{\beta S_0 p}{c\delta(1 + \frac{\mu}{\alpha})} \tag{S3}$$

.

The effective reproduction number at time of infection, R, from the immune model is less than or equal to R_0 depending on the initial level of immune effectors (Equation S2). For a fixed initial immune effector level, E_0, an approximation of R can be calculated by updating the δ term in R_0 with $\delta + kE_0$. While we assume $E_0 = 0$, the immune effector population does increase during the simulation so we calculated a time-dependent approximation of effective reproduction number, $R(t)$, by adjusting for both the current immune effector population and the remaining susceptible cell population:

$$R = R(t) = \frac{S(t)}{S_0} \frac{\delta}{\delta + kE(t)} R_0 \tag{S4}$$

.

Fitting the models to data.

Because these parameters were non-identifiable fitting to viral DNA concentration data alone, we fixed parameters using the literature (Supplementary Tables S1 - S2). Only β and the initial infection time (t_0) were fitted to the data with the first model (Equation S1). The model simulation begins with one latently infected cell ($L_0 = 1$ and $V_0 = 0$), and so t_0 represented how long the model was simulated before it was matched to the data at the first positive viral DNA concentration. This step was taken because the actual onset time of infection was not

observed given our weekly sampling interval. Squared error was then calculated comparing the viral DNA concentration from the model to the viral DNA concentration at matching swab times. Parameters values were then optimized by minimizing the sum of squares. A model was fit to each infant's data independently.

We estimated the parameters for R_0 (Equation S3) by fitting the model (Equation S1) only to the expansion phase of the data. We did this for three reasons: 1) we found poor performance of the target-cell model when fitting the entire data under our parameter assumptions; 2) we assumed that the immune system did not impact dynamics during the exponential expansion; and 3) we assumed the data best inform β and t_0 during this exponential expansion phase. We fit the immune model parameters to the entire episode using fixed β and t_0 from this analysis.

Because initial value selection may affect the search algorithm, we conducted 50 optimizations for each fit where the initial values for fitted parameters of the given model were selected using a Latin hypercube sampling scheme. The sampling constraints of the parameters were assumed to have very wide ranges, β : $10^{-14} - 10^{-6.5}$; t_0: $0.1 - 25$ days; θ: $10^{-3} - 100$; K_I: $1 - 10^{10}$; and γ: $10^{-5} - 1$. The optimization with the smallest sum of squared error was used.

Sensitivity analysis of target-cell limitation model.

Because the S compartment of the target-cell limitation model (Equation S1) is an abstraction of the susceptible cells in the oral cavity, the parameterization of this compartment contains uncertainty with regards to the growth and turnover (λ and μ) of cells. To explore if there were reasonable parameter ranges that could capture the kinetics of primary infections, we fit the model over several iterations of different assumptions regarding the death rate (μ) of these cells.

First, we made the assumption that, absent infection, that S cells were in a steady state

(a solution to the differential equations in Equation S1 where $V = 0$) such that

$$\lambda = \mu * S_0$$

Values for S_0 and μ are derived from a study looking at epithelial cell turnover (4). However, because CMV is known to infect other components of the oral cavity (specifically the parotid salivary gland (5)), these parameter values may be misspecified. The total amount of cells is unknown, but including the salivary glands, we assume that total cells (already set at 4×10^8) would not increase more than a log and choose to focus on the sensitivity of the μ parameter.

We chose seven levels of μ to explore over a large range (all per day): 1×10^{-4}, 1×10^{-3}, $\frac{1}{365}$, $\frac{1}{30}$, $\frac{1}{4.5}$, 1, and 2. Note that 1/4.5 is the value fixed from (4). For each fixed μ, we fit the model for each primary episode of CMV optimizing β, δ, and V_0 with c and p fixed (Table S1).

References

[1] Emery VC, Hassan-Walker AF, Burroughs AK, Griffiths PD. Human cytomegalovirus (HCMV) replication dynamics in HCMV-naive and -experienced immunocompromised hosts. J Infect Dis. 2002;185(12):1723–8.

[2] Stafford MA, Corey L, Cao Y, Daar ES, Ho DD, Perelson AS. Modeling plasma virus concentration during primary HIV infection. J Theor Biol. 2000;203(3):285–301.

[3] Moss P, Khan N. CD8(+) T-cell immunity to cytomegalovirus. Hum Immunol. 2004;65(5):456–64.

[4] Dawes C. Estimates, from salivary analyses, of the turnover time of the oral mucosal epithelium in humans and the number of bacteria in an edentulous mouth. Arch Oral Biol. 2003;48(5):329–36.

[5] Ho M. Pathology of Cytomegalovirus Infection. In: Ho M, editor. Cytomegalovirus: Biology and Infection. second edition ed.; 1991. p. 189–204.

[6] Heider JA, Bresnahan WA, Shenk TE. Construction of a rationally designed human cytomegalovirus variant encoding a temperature-sensitive immediate-early 2 protein. Proc Natl Acad Sci U S A. 2002;99(5):3141–6.

[7] Emery VC, Cope AV, Bowen EF, Gor D, Griffiths PD. The dynamics of human cytomegalovirus replication in vivo. J Exp Med. 1999;190(2):177–82.

[8] Shapiro M, Duca KA, Lee K, Delgado-Eckert E, Hawkins J, Jarrah AS, et al. A virtual look at Epstein-Barr virus infection: simulation mechanism. J Theor Biol. 2008;252(4):633–48.

Supplementary Tables

Table S1: Target-cell limitation model parameters

Parameter	Interpretation	Value	Notes	Ref.
S_0	Initial susceptible cell population	4×10^8	Fixed to total epithelial cell populations	(4)
L_0	Initial latently infected cell population	1	Start at first infected cell	
I_0	Initial infected cell population	0		
V_0	Initial viral population	0		
t_0	Days before first positive (infection day)	Fitted	Using expansion data only	
λ	Growth rate of susceptible cells per day	μS_0	Fixed assuming steady-state turnover over S	(4)
μ	Death rate of susceptible cells per day	1/4.5	Fixed assuming steady-state turnover over S	(4)
β	Rate of new infected cell per virus per susceptible cell per day	Fitted	Using expansion data only	
α	Lag (days) between cell infection and viral growth	1		(1; 6)
δ	Death rate of infected cells per day	0.77		(1; 7)
p	New viruses per infected cell per day	1600		(6)
c	Natural decay of virus per day	2	Used estimate from EBV oral clearance	(8)

Table S2: Immune model parameters

Parameter	Interpretation	Value	Notes	Ref.
S_0, L_0, I_0, V_0	Initial populations	Table S1		
$\lambda, \mu, \alpha, \delta, p, c$	Fixed infection parameters	Table S1		
t_0, β	Previously fitted infection parameters	Fitted previously	Fit from expansion phase analysis	
E_0	Initial immune effector population	0		
k	Immune removal rate per infected cell per day	0.01		(1; 2)
θ	Maximum growth of immune effectors per day	Fitted		
K_I	Infected cell level where growth of immune effectors 50% maximal	Fitted		
γ	Loss rate of immune effectors per day	Fitted		

Table S3: Mean differences among episode features comparing infants with HIV infected mothers to infants without HIV infected mothers.

Outcome	Estimate (95% CI)	P-value
Infection age (days)	-1.29 (-57.32, 54.75)	0.96
Peak ($\log_{10}$ CMV DNA copies/mL)	0.19 (-0.57, 0.96)	0.59
Expansion duration (days)	-6.86 (-57.19, 43.48)	0.77
Transition duration (days)	3.71 (-30.55, 37.98)	0.82
Clearance duration (days)	-97 (-259.87, 65.87)	0.22

Table S4: Bootstrap estimated Spearman's rank correlation coefficients between infection age (days) and episode features. * denotes significant at 95% significance level.

Outcome	Estimate (95% CI) (All infants)	Estimate (95% CI) (Excluding two older infants)
Peak ($\log_{10}$ CMV DNA copies/mL)	-0.59 (-0.85, -0.12)*	-0.38 (-0.76, 0.17)
Expansion duration (days)	0.2 (-0.52, 0.72)	-0.29 (-0.83, 0.43)
Transition duration (days)	-0.3 (-0.81, 0.31)	-0.16 (-0.69, 0.51)
Clearance duration (days)	0.05 (-0.55, 0.67)	0.19 (-0.47, 0.85)

Table S5: Results from cubic regression of clearance phase

Infant ID	Deceleration	Rebound	Inflection time (days)	Minimum exist?	Estimated minimum day	Minimum-peak	Final obs. day
M	*	*	159	TRUE	232	168	252
N	*	*	288	TRUE	241	59	290
F	*	*	73	TRUE	268	184	288
I	*	*	230	TRUE	317	191	357
J	*		237	TRUE	331	202	329
K	*		289	TRUE	436	347	320
C	*	*	222	TRUE	444	297	550
D	*		116	TRUE	481	418	481
A			502	FALSE			756
B			448	FALSE			572
E			98	FALSE			245
G			217	FALSE			386
H	*		234	FALSE			244
L	*		262	FALSE			295

Supplementary Figures

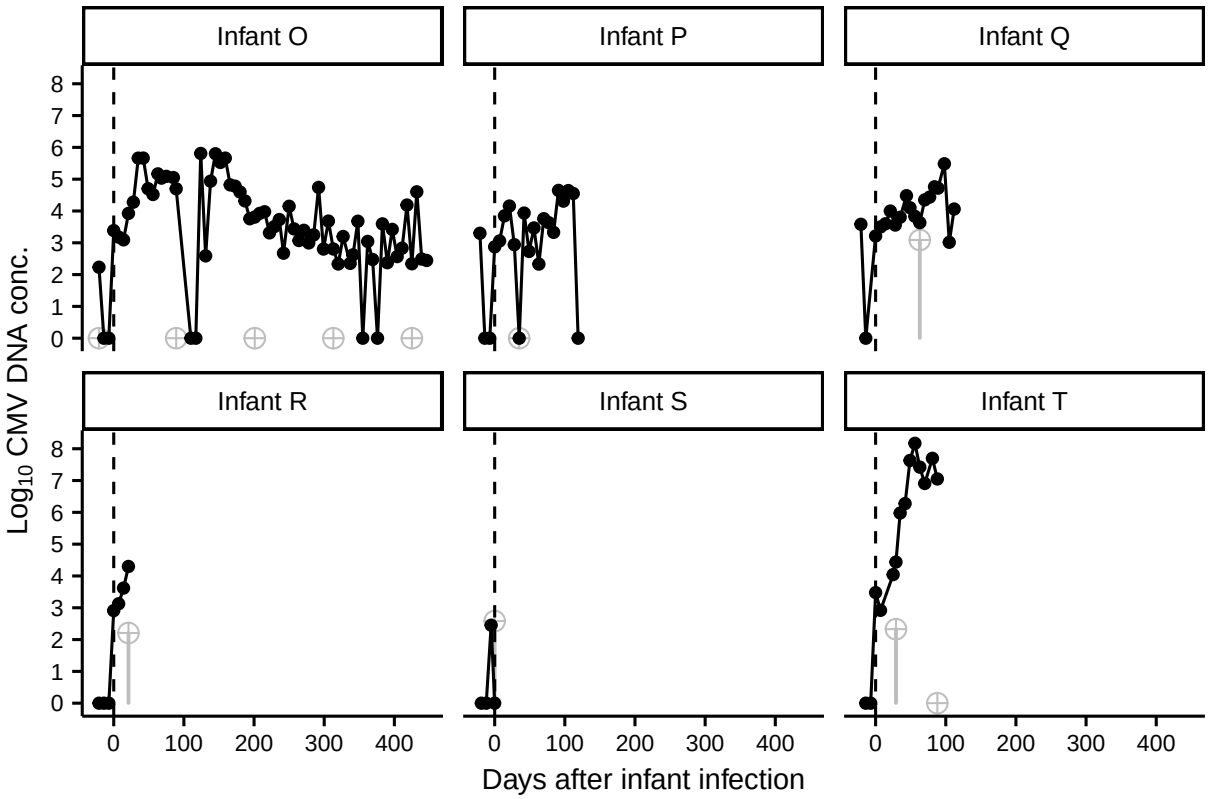

Figure S1: Oral shedding of CMV during primary infection for six excluded infants. Vertical gray bars with with fill circle denote time and magnitude of plasma CMV viral load. Plot starts 21 days before estimated infection time (denoted at 0 with vertical dotted line).

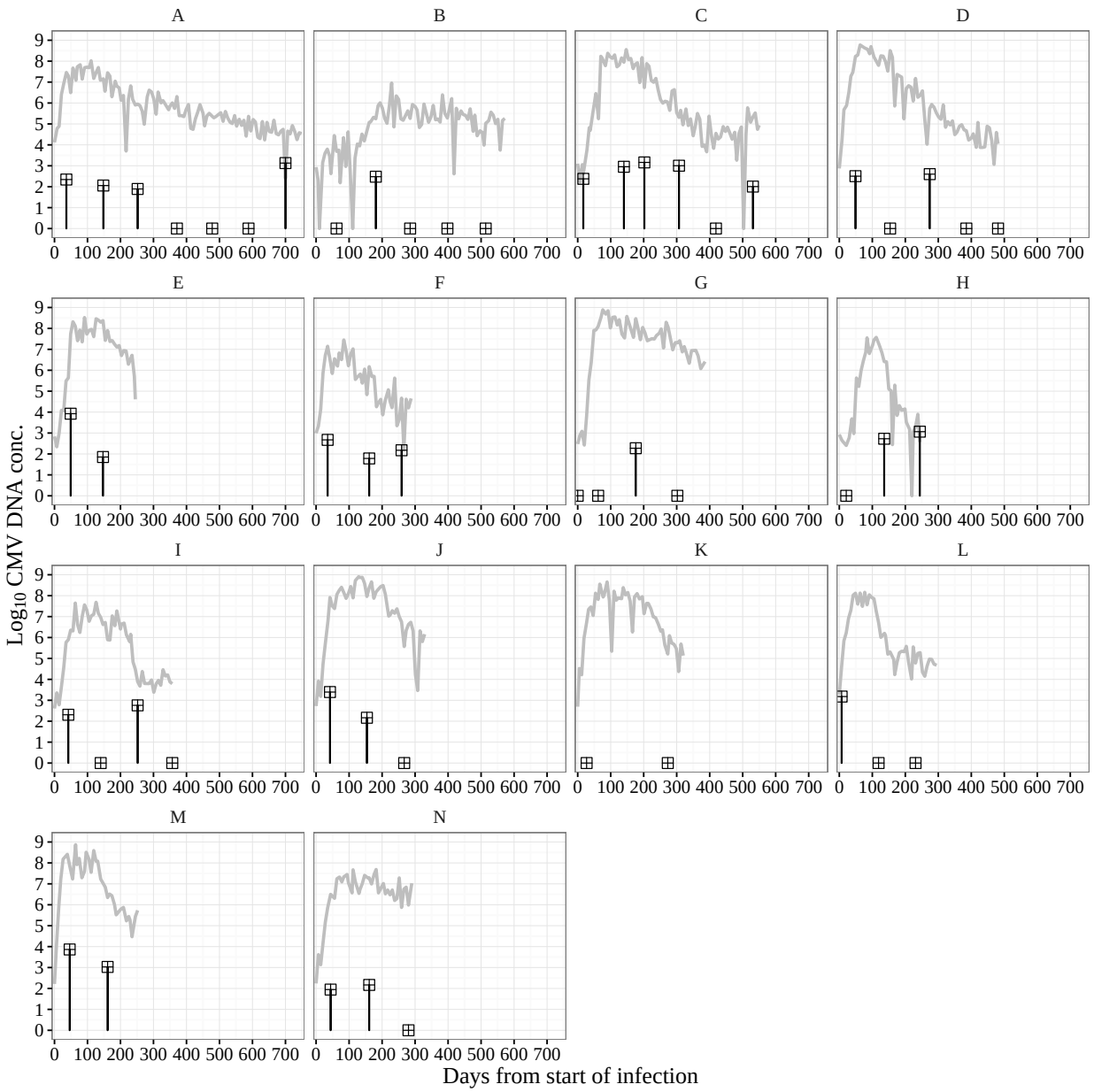

Figure S2: Oral and plasma CMV viral loads during primary infection for the 14 infants included in the main analysis. Vertical black bars with the filled square denote time and magnitude of plasma CMV viral load and gray tracings correspond to oral viral loads.

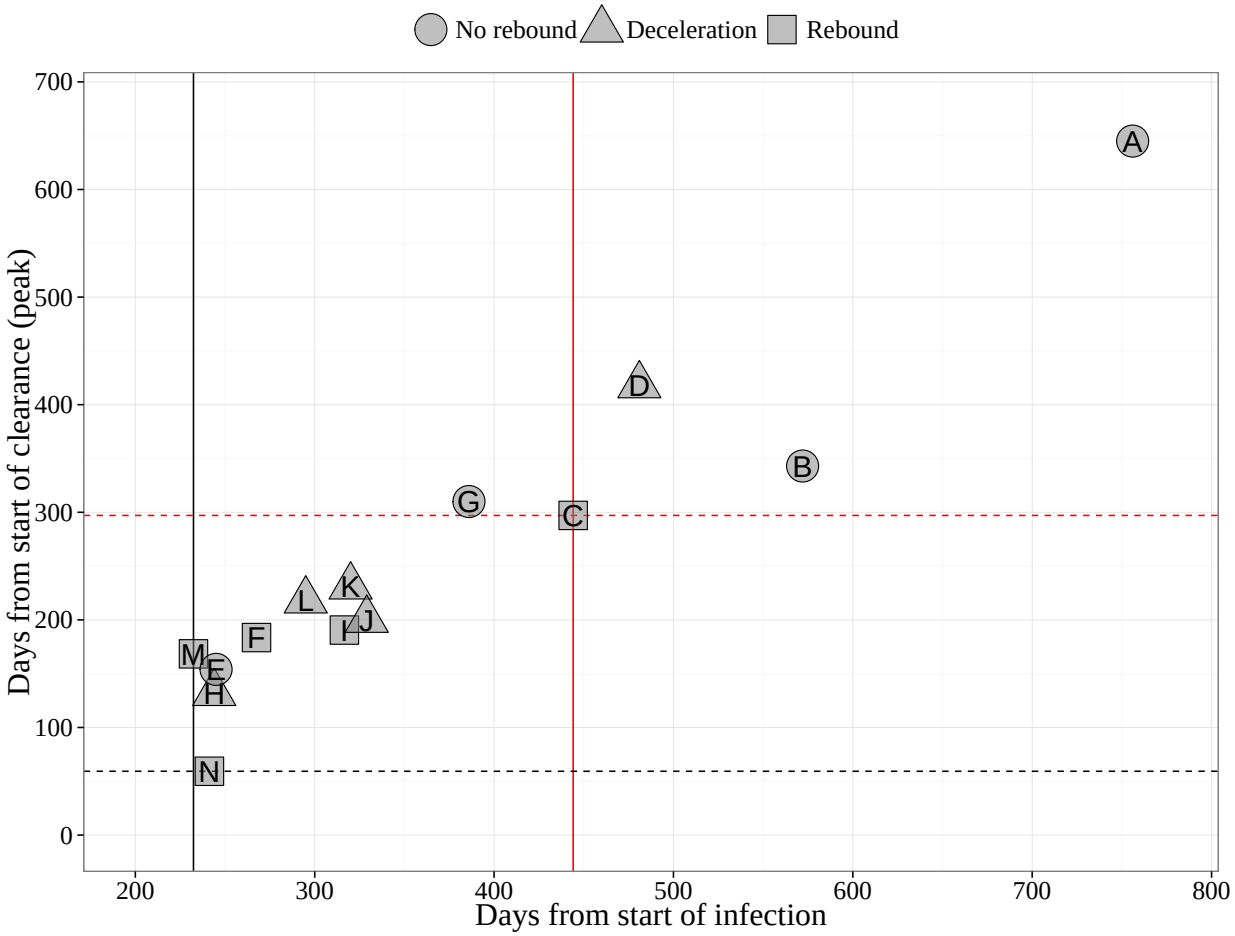

Figure S3: Timing of CMV shedding behavior during the clearance phase is highly variable. Points indicate the days from the start of infection and the end of the high shedding phase where either 1) rebounding was estimated (squares) or 2) the final observation day was observed (circles and triangles). Triangles represent infants where deceleration was observed but not viral rebound. The solid lines are the minimum (black) and maximum (red) days that are observed after infection starts before an estimated rebound day. The dotted lines are the minimum (black) and maximum (red) days that are observed after the end of the clearance phase (see text and Figure 1) before an estimated rebound day. Infant IDs are indicated by the letter inside each point.

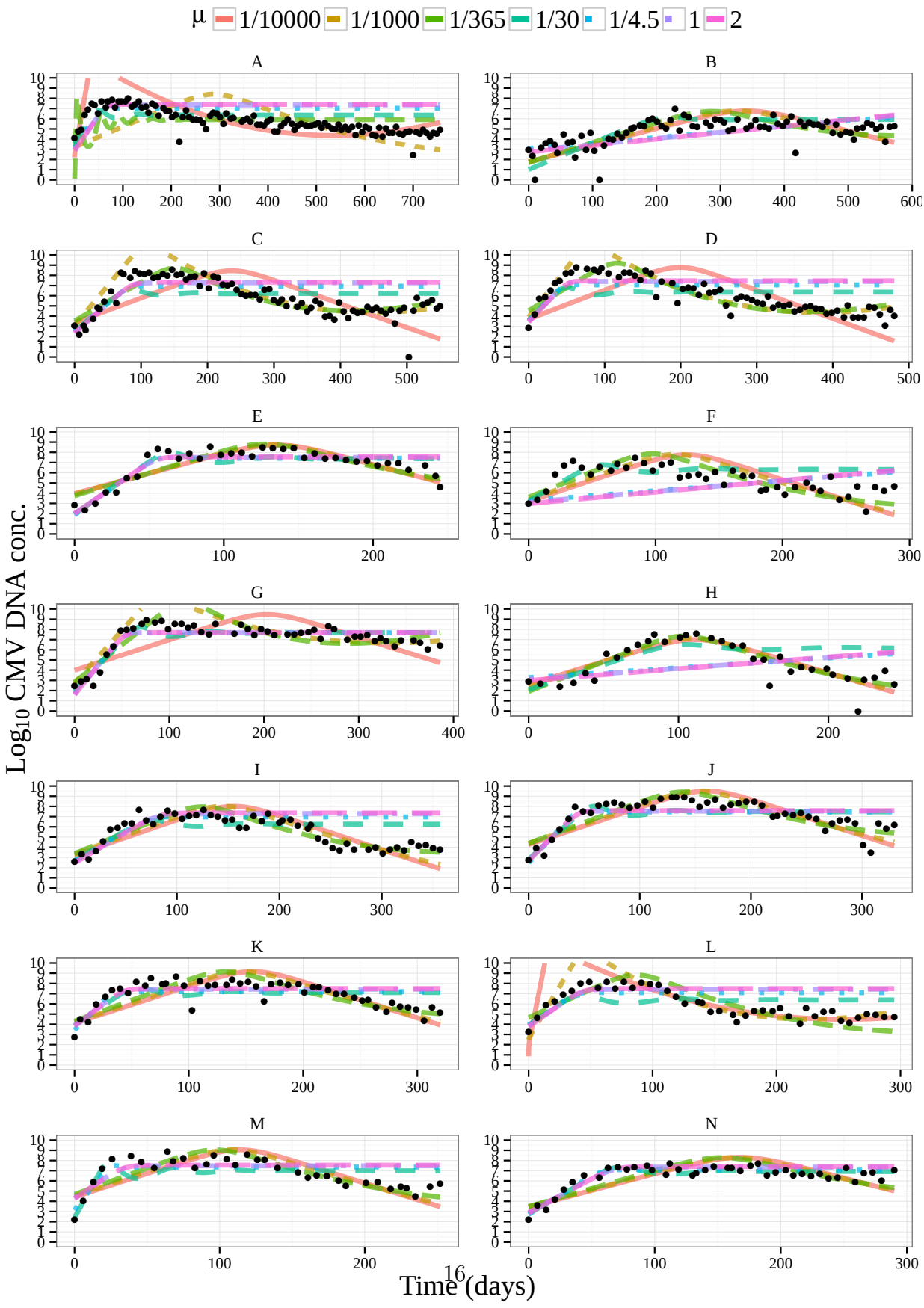

Figure S4: Target cell limitation model sensitivity fits over varying μ and optimized δ.

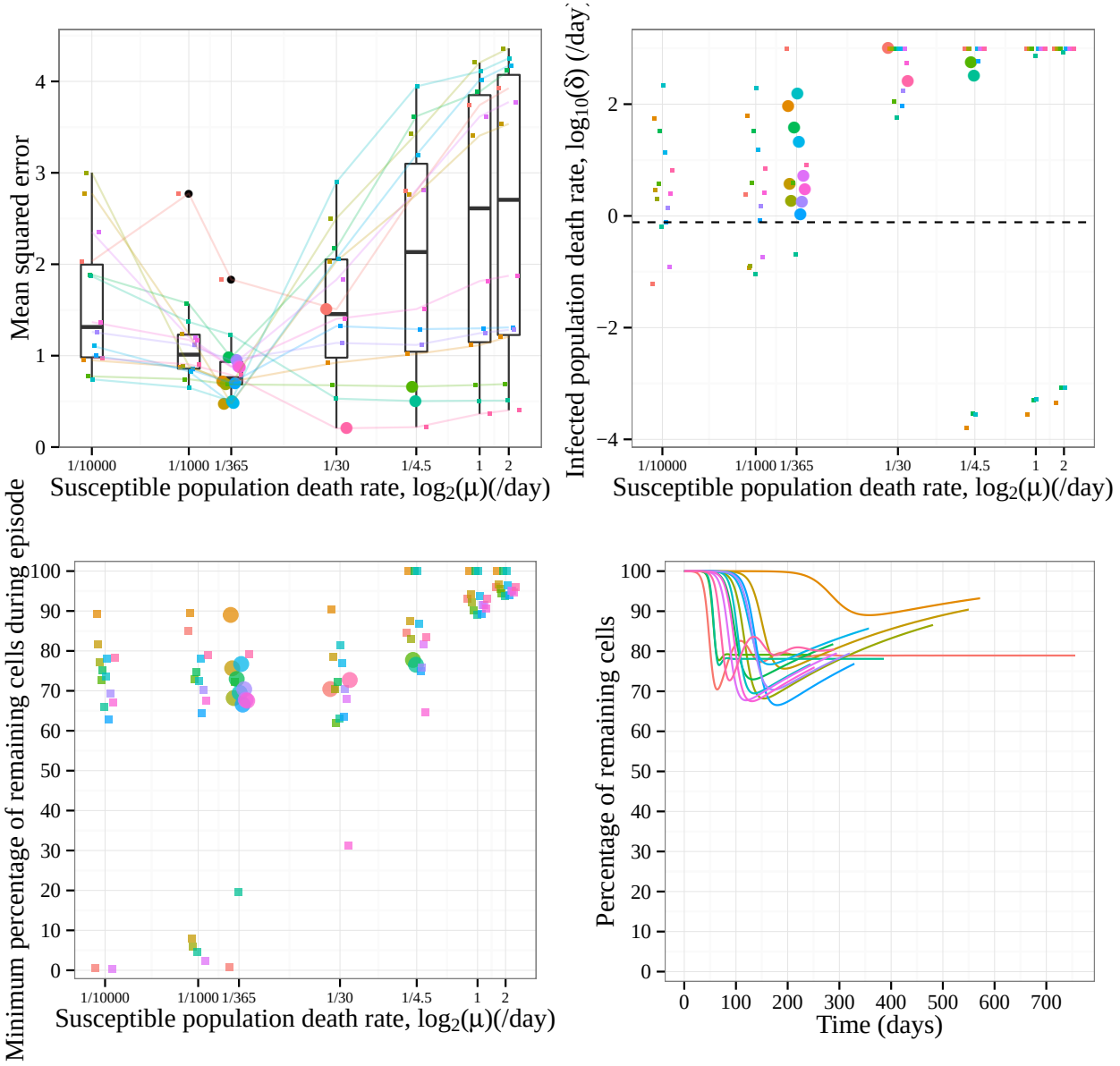

Figure S5: Optimization results for target cell model fits over varying μ and free δ. Each patient is denoted by different color. Smaller squares denote best fit for a given μ and larger circles denote the overall best fit for each patient (14 total). A) Best fit mean squared error for each μ value. B) Best fits δ for each μ value. The value used for δ in the main analysis is denoted by the horizontal dotted line. C) For each fixed μ, the minimum total cells remaining relative to the initial population (4×10^8) during the model simulation. For example, 10% indicates that at some point during the model simulation, only 10% of the total cells remained. D) The simulated dynamics of the total cell population relative to the initial population (4×10^8) using the best overall fit for each patient. The minimum of each curve corresponds to the large, similarly shaded circle in A-C. Model simulation time was limited to episode observation time for a given patient.